University of Warsaw
Faculty of Economic Sciences

Working Papers

No. 1/2010 (24)

Natalia Nehrebecka
Sylwia Grudkowska

Application of epsilon method to modeling expectations in construction

Warsaw 2010

Working Paper No. 1/2010 (24)

Application of epsilon method to modeling expectations in construction

Natalia Nehrebecka

Faculty of Economic Sciences
University of Warsaw

Department of Statistics
National Bank of Poland

Sylwia Grudkowska

Department of Statistics
National Bank of Poland

Abstract

The epsilon method has been applied to examine the strength of relations among selected objective and subjective factors connected with a manager's predictions of their companies' development. The aim of this research was to study which variable has the strongest impact on business expectations in the construction industry. The results offer compelling evidence that respondents rely both on their current opinion on enterprise as well as on general economic situation. The survey was carried out based on Polish data from 2000:1 to 2008:10.

Keywords:

Epsilon method, Multivariate Statistical Analysis, Singular Value Decomposition

JEL:

C1, M31

Introduction

The primary aims of the business tendency survey are forecasting economic activity, identifying current economic condition of the country and predicting the direction of changes in business activity. Business survey indicates are a valuable tool for economic analysis for all types of users. Moreover, this type of survey provides information on assessments and expectations of the economic situation by actors on the market. Conventional methods often prove insufficient in collection of information regarding business tendency surveys. Therefore, it is a useful source of information about changes in aggregated economic activity.

Business tendency survey in construction usually includes: capacity utilization, plans and expectations for immediate future, managers' views on the overall economic situation[1]. The selective key survey variables are used to construct a confidence indicator which reflects the general business tendency.

This article aims at explaining the mechanism of formulation the respondent's expectation about business tendency in construction. It pursues factors which are closely associated with subjective expectations about future company's development. It investigates whether indices of current company's sentiment have a strong impact on future business tendency indicators. On the other hand, this paper also includes the objective, macroeconomic data, like unemployment or gross wages. The purpose is to check if the signal from real economy strongly influences managers' predictions of development of their companies. The leading hypothesis of the article is that respondents rely on their own evaluation rather than on the general economic situation.

Comparison between these two groups of regressors will be based on a relative weight method. We apply this method to estimate the relative importance of the predictor variable. The method and its usefulness will be presented in the next section. Section 2 discusses the data used in the empirical analysis. Section 3 reports the results which we obtained for the business tendency indicators. Finally, section 4 provides a conclusion.

[1] Business Tendency Surveys: A Handbook, OECD

1. The epsilon method

Econometric modeling is facing a serious problem because of correlation between explanatory variables. The reason why the correlation occurs is interdependence of economic phenomena. In case of strong relationship between independent variables it is hard to separate the particular variable's influence on dependent variable from other variables. This problem makes economic analysis really complicated. It was the reason to search for a measure of relative importance of variables in multiple regression. The history of this research was described by Johnson and LeBreton (2004)[2]. Each of methods mentioned (zero - order correlations, standardized regression coefficients, not standardized regression coefficients, semi-partial correlation, the product measure, average squared semi-partial correlation, average squared partial correlation) has its drawbacks which makes this technique useless and leads to wrong conclusions.

These studies resulted in two new important methodologies - dominance analysis and epsilon method. First of these was introduced by Budescu (2003)[3]. Dominance analysis allows ranking the predictor variables. It is said that the variable A is dominate to variable B if A it is more useful than B in all subset multiple regressions. Given that the above definition is really strict, it is difficult to implement it for data sets of significant size, because it requires estimating $2^p - 1$ submodels, where p is the number of variables in the original model. The advantage of this method is that relative weights obtained from dominance sum to the model R^2. Because of that relative importance of each variable can be calculated as the proportion of predictable variance for which it accounts (Johnson, LeBreton (2004))[4].

This article applies the second approach - the epsilon method. It is much easier to implement that this dominance analysis and lead to similar results[5].

The epsilon procedure presented below derives from Johnson (2000). The first step involves the following decomposition of predictor scores matrix:

---

[2] Johnson, J., LeBreton J. (2004), "History and Use of Relative Importance Indices in Organizational Research",
[3] Budescu D. (1993), "Dominance analysis: a new approach to the problem of relative importance of predictors in multiple regression"
[4] Johnson, J., LeBreton J. (2004), "History and Use of Relative Importance Indices in Organizational Research",
[5] Johnson J. (2000), "A Heuristic Method for Estimating the Relative Weight of Predictor Variables in Multiple Regression"

$$\mathbf{X} = \mathbf{P}\mathbf{\Delta}\mathbf{Q}^T \tag{1}$$

where:

$\mathbf{X}$ is matrix *n x p*,

$\mathbf{P}$ is the full matrix of eigenvectors of $\mathbf{X}\mathbf{X}^\mathbf{T}$,

$\mathbf{Q}$ is the full matrix of eigenvectors of $\mathbf{X}^\mathbf{T}\mathbf{X}$,

T denotes transposition,

$\mathbf{\Delta}$ is a diagonal matrix which contain the square roots of the eigenvalues of ***XX'*** and ***X'X***.

If none of the vectors from the matrix $\mathbf{X}$ is collinear with another one, each of element of ***XX'*** and ***X'X*** is different from zero. In this case the best - fitting approximation of $\mathbf{X}$ is:

$$\mathbf{Z} = \mathbf{P}\mathbf{Q}^T \tag{2}$$

The above equation ensures that the residual sum of squares between the raw predictors and the orthogonal variables is minimized. The matrix $\mathbf{Z}$ contains z_k variables which are orthogonal to the original predictors x_j from the matrix $\mathbf{X}$. The matrix $\mathbf{Z}$ is then used to obtain the set of relative weights. The calculation is based on regression $\mathbf{X}$ on $\mathbf{Z}$:

$$\Lambda^* = (\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{X} = (\mathbf{Q}\mathbf{P}^T\mathbf{P}\mathbf{Q}^T)^{-1}\mathbf{Q}\mathbf{P}^T\mathbf{P}\mathbf{\Delta}\mathbf{Q}^T = \mathbf{I}^{-1}\mathbf{Q}\mathbf{\Delta}\mathbf{Q}^T = \mathbf{Q}\mathbf{\Delta}\mathbf{Q}^T \tag{3}$$

As the Z_k are not correlated, the correlation coefficients λ^*_{jk} are equal to the correlation between x_j and z_k. The squared elements of Λ^* defined as λ^*_{jk} are equal to the proportion of predictable variance in x_j accounted for z_k. The relative contribution x_j to y can be approximated by $\sum_k \lambda^{*2}_{jk}\beta^{*2}_k$, where β^*_k is regression coefficient taken from least square estimation of y to z_k.

The coefficient vector β^* is calculated as:

$$\beta^* = (\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{y} = (\mathbf{Q}\mathbf{P}^T\mathbf{P}\mathbf{Q}^T)^{-1}\mathbf{Q}\mathbf{P}^T\mathbf{y} = \mathbf{I}^{-1}\mathbf{Q}\mathbf{P}^T\mathbf{y} = \mathbf{Q}\mathbf{P}^T\mathbf{y} \tag{4}$$

Each z_k is a linear combination of the x_j. The variance in y accounted for by z_k is, therefore, partitioned among the according to the proportion of variance in z_k accounted for by each x_j.

In this algorithm one can obtain the set of relative weights which sum to the model R^2. Each weight is attributed to the particular raw independent variable and reflects the effect of the variable by itself and in combination which the others.

2. The data

The sample period for all results is 2000:1- 2008:10. Data have been provided by the Polish Central Statistical Office. The dependent variables are equal to the following construction expectation: expected general economic situation of the enterprise, expected prices of construction/assembly works carried out by the enterprise, expected employment in construction/assembly operations of the enterprise. These indicators are equal to the difference between weighted percentages of positive (the first option) and negative answers (the third option). Weighting is necessary as the companies which are taking part in the survey vary in size. The weights are based on the value of sold construction assembly production in current process[6]. The questions are presented in table 1.

TABLE 1. Business Tendency questionnaire

Question	Possible answers
Expected financial situation of the enterprise	• will be better • will remain unchanged • will be worse
Expected prices of construction/assembly works carried out by the enterprise	• will be better • will remain unchanged • will be worse
Expected employment in construction/assembly operations of the enterprise	• will be better • will remain unchanged • will be worse

Source: Monthly questionnaire in construction, Central Statistical Office, www.stat.gov.pl.

[6] http://www.stat.gov.pl/cps/rde/xbcr/gus/PUBL_business_tendency_survey_in_manufacturing_2.pdf

Each year, ca. 2200 construction enterprises of more than 9 employees take part in the survey. The set of independent variables which could influence expectations consist of two parts. In the first one there are business tendency indicators of current economic situation in construction. This group includes:

- Construction/assembly works carried out by the enterprise at domestic market,
- The guaranteed operating period of the enterprise,
- Capacity utilization of the enterprise[7],

The second group consists of macroeconomic indicators:

- Index of construction and assembly production in constant prices (2000=100),
- Average monthly gross wages excluding payments from profit in construction,
- Dismissals declared by enterprises[8],
- Average employment in construction,
- Housing loans of Monetary Financial Institution[9] to households,
- Unemployment rate - registered[10],
- Dwellings under construction,
- Sale of construction and assembly production,
- Dwellings completed.

These data derived from polish Central Statistical Office and National Bank of Poland. The time series used in our analysis were seasonally adjusted using TRAMO/SEATS procedure[11]. It was done due to the fact that in each series the seasonality was detected and in most cases the seasonal component was strong. In this situation using raw data could result in spurious correlation between variables.

[7] in percentages.
[8] Number of persons which will be dismissed declared by enterprises.
[9] Excluding National Bank of Poland.
[10] Calculation based on number of unemployed registered in employment offices.
[11] Maravall, A. (2005), "An Application of the automatic procedure of TRAMO and SEATS; Direct versus Indirect Adjustment".

3. Results

Table 2 presents relative weights calculated for our dependent variables. These weights were rescaled by dividing them by the model R^2 and multiplying by 100, so they are equal to percentage of the model R^2 connected with each predictor.

TABLE 2. Relative weights

Variable	Expected financial situation of the enterprise	Expected prices of construction/assembly works carried out by the enterprise	Expected employment in construction/assembly operations of the enterprise
Construction/assembly works carried out by the enterprise	17,60%	14,70%	15,70%
Capacity utilization of the enterprise	2,30%	2,50%	2,40%
The guaranteed operating period of the enterprise	12,60%	12,70%	11,40%
Housing loans of MFI's to households	7,00%	7,30%	8,20%
Dwellings completed	0,50%	0,30%	0,70%
Dwellings under construction	8,00%	4,70%	9,20%
Average gross wages excl. payments from profit in construction	7,20%	14,30%	6,90%
Unemployment rate - registered	6,30%	6,90%	7,70%
Dismissals declared	10,10%	10,40%	8,30%
Average paid employment in construction	7,30%	7,30%	9,40%
Sale of construction and assembly production	8,60%	9,60%	7,30%
Index of construction and assembly production	12,50%	9,40%	12,90%

Source: own calculations

Expectations regarding financial situation

Results presented in Table 2 indicate that construction/assembly works carried out by the enterprise was accounted for almost 18 of predictable variable of expected financial situation of the enterprise. It is because this indicator denotes the number of orders which will be carried out. The guaranteed operating period of the enterprise has also quite a share in R^2 (12,6%) which only supports our feeling that current orders has a dominant role in predicting the financial situation. The longer the guaranteed operating period of the enterprise is the more optimism about chances of financial success the respondents' display.

The current situation in construction sector (index of construction and assembly production) as well as sold construction and assembly production is also a valuable predictor of the future financial situation.

It is noticeable that data connected with employment are moderately important for respondents. For example, average monthly gross wages excluding payments from profit in construction which accounts for more than 7% of R^2. It is a direct result of rapid growth in construction during past few years in Poland. In this period the wages in construction went up because of labor shortage. The moderate relative weight of unemployment rate and dismissals declared indicates that respondents follow the situation on the labor market, particularly the future tendency (the relative weight of dismissals declared is higher than the relative weight of unemployment rate).

The number of buildings completed plays no role in explaining the variation of dependent variable, due to the fact that it takes quite a long time to build a new building. On the contrary, respondents take into account the number of dwellings under construction as it is closely related with future price of dwellings. Value of housing loans of MFI's to households plays a similar role. This variable accounts for 7% of R^2 and indicates the demand for new housing (the majority of buyers purchase property on the primary market).

Surprisingly, the relative weight of utilization capacity was below 2%. However, it is possible that enterprise could easily increase or decrease their utilization capacity to adjust it to the current needs. If so, this factor is not essential in explaining the dependent variables' variance.

Expectations regarding prices in construction

The result of our survey shows that the expectations regarding prices in construction are similar to the outcome of the previous model. It is not a surprise that the average monthly gross wages excluding payments from profit in construction has almost the highest relative weight in this model. As the level of average wages is rising the company's profit is reduced. The company is unable to avoid these expenses, because it needs to fulfill contractual obligations. The other labor market's variables - unemployment rate and dismissals declared - have even higher influence on expectations regarding prices of construction than expectation regarding financial

situation of the enterprise. These variables have a strong influence on the level of wage. Obviously, the wage claims are directly reflected in the prices of construction. On the other hand the influence on financial situation of the enterprise is not so strong - there are a set of other factors which also affect it.

Construction/assembly works carried out by the enterprise have a very similar wage as average monthly gross wages. It is because they are the fixed term contracts so provisions of an agreement cover the prices. On this basis the enterprise can estimate their future revenues and costs.

Expectations regarding employment in construction

In comparison to models discussed above, the average employment in construction has quite a strong influence on expectations of employment. It is because the high wage level reduces the enterprise's capability to increase employment.

It is not strange that the number of dwellings under construction and the housing loans of MFI's to households occur to be a satisfactory predictor of dependent variable. If dwellings are under construction it is obvious that enterprises need their employees to finish them, so they would not reduce the labor force in near future. The value of housing loans indicates that households are interested in property purchase. This trend has been observed in Poland since 2002, so respondents anticipate that it will remain unchanged.

4. Conclusions

For the purpose of this article we have applied the epsilon method in order to describe the weighting strategies used by respondents when making an overall performance evaluation. Although this procedure is valid when the dependent variables are correlated, it is recommended to consider that none of the two variables measures the same thing. Were it the case, one of them should be removed as the relative weight will spread evenly between them. Keeping both predictors could artificially inflate the weight of a construct that is measured by only one variable (Johnson, 2000).

Results of our survey show that both the enterprise's evaluation of the current economic activity and the general economic situation in Poland have a strong impact on respondent's expectation about business tendency in construction. The most valuable predictors are construction/assembly works carried out by the enterprise and the guaranteed operating period of the enterprise. The average relative weight was equal to 16,0% and 12,2% respectively. On the other hand, average monthly gross wages excluding payments from profit in construction, as well as the index of construction and assembly production, which are included in general economic situation's indicators, have influenced explaining the variance of dependent variables (13,7% and 11,4% respectively).

References

Budescu D. (1993), "Dominance analysis: a new approach to the problem of relative importance of predictors in multiple regression", *Psychological Bulletin*, no. 114

Business Tendency Surveys: A Handbook, OECD

Budżety gospodarstw domowych w 2007 r., Główny Urząd Statystyczny, Warszawa 2008 r.

Dominitz J., Manski C. (1994), "Using Expectations Data to Study Subjective Income Expectations", National Bureau of Economic Research

Forsells M., Kenny G. (2002), "The rationality of consumers' inflation expectations: survey based evidence for euro area", *European Central Bank Working Papers*, n. 163, Frankfurt

Johnson, J. (2004), "Factors Affecting Relative Weights: The influence of Sampling and Measurement Error", *Organizational Research Methods*, Vol. 7, no. 3, 283-299

Johnson, J., LeBreton J. (2004), "History and Use of Relative Importance Indices in Organizational Research", *Organizational Research Methods*, Jul 2004

Johnson J. (2000), "A Heuristic Method for Estimating the Relative Weight of Predictor Variables in Multiple Regression", *Multivariate Behavioral Research*, no. 35

LeBreton J., Ployhart R., Ladd R. (2004), "A Monte Carlo Comparison of Relative Importance Methodologies", *Organizational Research Methods*, Vol. 7, no. 3

Maravall A. (2005), "An Application of the automatic procedure of TRAMO and SEATS; Direct versus Indirect Adjustment", *Banco de España Working Papers* n. 0524

Wall M., Rechtsteiner A., Rocha L. (2003), "Singular value decomposition and principal component analysis", in: Berrar D., Dubitzky W., Granzow M., "A Practical Approach to Microarray Data Analysis", Springer

Westerhoff F. (2006) „Nonlinear expectation formation, endogenous business cycles and stylized facts", *Studies in Nonlinear Dynamics and Econometrics*, Vol. 10

Faculty of Economic Sciences
University of Warsaw
44/50 Długa St.
00-241 Warsaw
www.wne.uw.edu.pl